PCT

[redacted]D INTELLECTUAL PROPERTY ORGANIZ[redacted]
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :		(11) International Publication Number:	WO 93/02198
C12N 15/62, 15/13, C07K 13/00 // A61K 39/395	A1	(43) International Publication Date:	4 February 1993 (04.02.93)

(21) International Application Number: PCT/FI91/00225

(22) International Filing Date: 23 July 1991 (23.07.91)

(71) Applicant *(for all designated States except US)*: TECHNICAL RESEARCH CENTRE OF FINLAND [FI/FI]; Vuorimiehentie 5, SF-02150 Espoo (FI).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)* : TEERI, Tuula, T. [FI/FI]; Porttitie 17B, SF-02180 Espoo (FI). TAKKINEN, Kristiina [FI/FI]; Haltilantie 12 as. 1, SF-02200 Espoo (FI). LAUKKANEN, Marja-Leena [FI/FI]; Ahvenevä 1B16, SF-02170 Espoo (FI). ALFTHAN, Kaija [FI/FI]; Isonnevantie 1A11, SF-00300 Helsinki (FI). SIZMANN, Dorothea [DE/FI]; Salakkatie 7, SF-02170 Espoo (FI). KNOWLES, Jonathan, K., C. [FI/FI]; Ritokalliontie 15A, SF-00330 Helsinki (FI).

(74) Agent: OY KOLSTER AB; Iso Roobertinkatu 23, P.O. Box 148, SF-00121 Helsinki (FI).

(81) Designated States: AT, AU, BB, BG, BR, CA, CH, CS, DE, DK, ES, GB, HU, JP, KP, KR, LK, LU, MC, MG, MN, MW, NL, NO, PL, RO, SD, SE, SU, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LU, NL, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, SN, TD, TG).

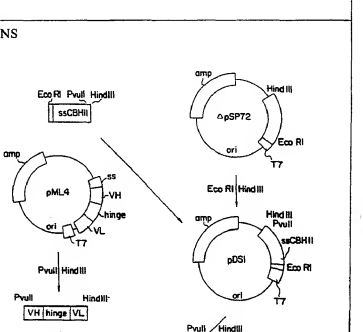
Published

With international search report.

(54) Title: RECOMBINANT SECRETABLE FUSION PROTEINS

(57) Abstract

A method for producing a single chain fusion protein, comprises the steps of: (a) constructing an rDNA molecule coding for a secretable single chain fusion protein having a plurality of functional proteins or protein domains joined by spacer peptides, wherein each such spacer peptide comprises a flexible hinge region of a naturally secreted multidomain protein or an analogue thereof, each adjacent pair of functional proteins or protein domains is joined by a spacer peptide which is not naturally continuous with either functional protein or protein domain to which it is joined, and at least two of the functional proteins or protein domains each independently exhibit physicochemical or biological activity or interact to form an assembly that exhibits physicochemical or biological activity; (b) inserting the rDNA molecule into an expression vector capable of transforming a host organism; (c) transforming a host organism with the resultant vector, and isolating transformants; and (d) culturing the transformants and recovering thesingle chain fusion protein.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**FI**	Finland	**ML**	Mali
AU	Australia	**FR**	France	**MN**	Mongolia
BB	Barbados	**GA**	Gabon	**MR**	Mauritania
BE	Belgium	**GB**	United Kingdom	**MW**	Malawi
BF	Burkina Faso	**GN**	Guinea	**NL**	Netherlands
BG	Bulgaria	**GR**	Greece	**NO**	Norway
BJ	Benin	**HU**	Hungary	**PL**	Poland
BR	Brazil	**IE**	Ireland	**RO**	Romania
CA	Canada	**IT**	Italy	**RU**	Russian Federation
CF	Central African Republic	**JP**	Japan	**SD**	Sudan
CG	Congo	**KP**	Democratic People's Republic of Korea	**SE**	Sweden
CH	Switzerland			**SN**	Senegal
CI	Côte d'Ivoire	**KR**	Republic of Korea	**SU**	Soviet Union
CM	Cameroon	**LI**	Liechtenstein	**TD**	Chad
CS	Czechoslovakia	**LK**	Sri Lanka	**TG**	Togo
DE	Germany	**LU**	Luxembourg	**US**	United States of America
DK	Denmark	**MC**	Monaco		
ES	Spain	**MG**	Madagascar		

RECOMBINANT SECRETABLE FUSION PROTEINS

Background of the Invention

The present invention relates to a method for producing secretable, biologically active single chain antibodies (scAbs) and other secretable fusion proteins having at least two distinct functional proteins or protein domains linked to form a single polypeptide chain by a flexible linker peptide that is compatible with secretion. Also disclosed are recombinant DNA (rDNA) molecules coding for the fusion proteins, expression vectors and transformed hosts suitable for use in the method, as well as fusion proteins having known or new functions which can be made by the foregoing method.

Proteins are the key elements in a wide variety of biotechnological applications today and their importance is expected to increase significantly in the future. Genetic and protein engineering are used widely to study and improve properties of proteins. The key techniques include site-directed and random mutagenesis of the isolated genes or cDNAs coding for the desired proteins and the design and construction of fusion proteins using genetic engineering. Increased understanding of protein properties and function will make possible the design and development of completely novel proteins with new activities.

Today, however, protein engineering and design are largely dependent on the building blocks of protein structures provided by nature. Fusion

proteins have been constructed by linking together two previously independent proteins or protein domains with varying degrees of success. Most commonly, an enzymatically detectable marker protein or an affinity tail facilitating protein purification has been linked to the N- or C-terminus of another protein. In most cases, the fusions have been accomplished by trial and error and only very rarely have attempts been made to design linkers or spacers for the optimization of the fusion point.

Immunoglobulins are large, complicated molecules consisting of two identical light chains and two identical heavy chains which pack together forming the complete immunoglobulin. Each of the heavy chains is composed of 4-5 domains and each of the light chains of 2 domains. Immunoglobulins are unique in their capacity to recognize and bind specifically a wide variety of molecular structures. This capability can be used in many applications for the detection or purification of compounds, for targeting of drugs or marker proteins to specific tissues and for other applications for which specific molecular recognition and binding is important.

In many biotechnical applications, e.g., diagnostics, only the specificity determining regions (hypervariable domains) of the complete immunoglobulin molecule are required. Fab fragments have been produced proteolytically from intact immunoglobulin or genetically using truncated immunoglobulin genes expressed in bacteria (Better et al., *Science*, 240:10410-1043, 1988; Cabilly et al., *Eur. Pat. Appln. No.* 0125023, published 11/14/84, corresponding to *U.S. Serial No.* 483,457, filed 4/8/83). There are

published methods for the secretion of the two domains of Fab and Fv fragments as separate peptides in E. coli (Skerra et al., *Science*, 240:1938-1041, 1988; Better et al., 1988, *supra*).

Single chain antibodies (scAbs), in which the antigen binding domains of an immunoglobulin have been linked to a single polypeptide chain using short flexible linkers or spacers have been constructed and produced in bacteria (Bird et al., *Science*, 242:423-426, 1988; Ladner et al., *PCT application* WO 88/01649, 1988; Huston et al., *Proc. Natl. Acad. Sci. USA*, 85:5879-5883, 1988; Chaudhary et al., *Nature*, 339:394-397, 1989). In all of the published cases, the single chain antibodies were produced inside of E. coli and required solubilization and subsequent renaturation to obtain active protein, even in cases where the linker peptides were designed for single chain antibodies using all available information of natural linker sequences (Bird et al., 1988, *supra*; Ladner et al., 1988, *supra*). However, for efficient and economically feasible exploitation of the engineered immunoglobulins their secretion to the culture medium of the production host is required.

In the literature there are few examples of linker sequences used in constructing novel fusion proteins. Three examples concern the construction of single chain antibodies using two different types of linker sequences. In two of the examples the linker is a thrice-repeated sequence of Ser-Gly-Gly-Gly-Gly, (Chaudhary et al., 1989, *supra*; Huston et al, 1988, *supra*) and in the third example an 18 residues long linker peptide was designed by computer aided modelling (Bird et al., 1988, supra; Ladner et al.,

1988, *supra*). As noted above, the proteins in each of these cases were produced inside of Escherichia coli cells and solubilization and renaturation were required in order to obtain active fusion protein.

Computer-aided molecular modeling programs used today for protein design utilize protein structural information contained in data banks such as the Brookhaven protein databank. The data collected in these data banks are obtained almost exclusively from protein structures solved by x-ray crystallography. Only the structures of very few small proteins (less than 20 KDa) have been solved by nuclear magnetic resonance (NMR), and these consist of only a single domain. In designing linker sequences suitable for scAbs, the data banks have been used as a course for generating the starting structures (Ladner et al., 1988, *supra*). However, it has been observed, e.g., that intact fungal cellulases are not readily crystallizable while the core proteins without the hinge and the substrate recognition domain can be crystallized (Bergfors et al., *J. Mol. Biol.*, 209:167-169, 1989).

In a recent publication, Argos describes the properties and sequences of a number of natural linker sequences found at the domain interfaces of natural proteins included in the Brookhaven protein structural databank (Argos, *J. Mol. Biol.*, 211:943-958, 1990). These interdomain linkers are short segments containing an average of 5-6 amino acid residues which are mainly small polar amino acid residues with occasional charged residues in between, but which are not large basic and acidic residues like Arg and Glu.

It was suggested that the best general linker consists of Ser, Thr and Gly residues only.

It is likely that longer hinges are not included since their intrinsic flexibility prevents crystallization. Therefore, important structural information concerning the architecture of flexible regions in proteins is likely to be missing from such data banks and for this reason, the Ladner et al. approach is not likely to be readily adapted to the design of linkers compatible with secretion.

In one recent example, polymers of human CuZn superoxide dismutase were constructed by linking the subunits together using the hinge peptide of an IgA1 (Hallewell et al., *J. Biol. Chem.*, 264:5260-5258, 1989). In this case, no signal sequence was included and the fusion proteins were produced inside E. coli or yeast cells and were not secreted.

Two recent papers disclose properties of the linker peptides found in bacterial multidomain regulatory proteins and multienzyme complexes (Wootton et al., *Protein Engineering*, 2:535-543, 1989; Radford et al., *J. Biol. Chem.*, 264:767-775, 1989). The Q-linkers described by Wootton et al. are a class of interdomain sequences found in bacterial multidomain regulatory proteins. They are approximately 20 amino acid residues long, relatively rich in Glu, Arg, Gln, Ser and Pro residues and found primarily in prokaryotic two-component regulatory and signal transduction systems. Radford discloses that the interdomain linkers of about 30 amino acids long from E. coli pyruvate dehydrogenase multienzyme complex consist mainly of Ala and Pro residues with occasional Gln, Glu, Lys, Arg, Ser or Thr residues.

The proteins used as sources for the linker peptides are intracellular and there are no published reports of the analysis or design of secretable linker peptides. Yet in industrial applications, secretion of the product is in most cases desired because the recovery of the product is easier and thus cheaper from the culture media than from inside of the producer organism. Perhaps more importantly, the proportion of correctly folded and thus biologically active proteins is generally very low (less than 10%) after solubilization and subsequent renaturation of the proteins produced intracellularly while the major fraction of secreted proteins are as active as the wild type protein.

Most if not all of the production organisms commonly used for expressing recombinant proteins are capable of secretion. These include many different bacteria, yeasts, fungi and plant and animal cells. A number of naturally secreted proteins have been cloned and successfully secreted from novel host organisms using signal sequences of naturally secreted proteins of the host or related organisms. However, there are no published examples to date of novel fusion proteins with added heterologous linker sequences which are secreted to the culture medium of the host. This is believed to be in part due to the fact that the rules for protein folding are still very poorly known and therefore it is generally not possible to predict the behavior of a protein or a part thereof based on its amino acid sequence.

Carbohydrate degrading enzymes, e.g., cellulases and glucoamylases are secreted proteins many of which consist of at least two domains; a large

catalytic core domain and a smaller substrate recognition domain (van Tilebeurgh et al., *FEBS Lett.*, 204:223-227, 1986; Teeri, Doctoral Thesis, Technical Research Centre of Finland, Publications 38, *Espoo,* 1987; Teeri et al., Gene, 51:43-52, 1987; Knowles et al., *Tibtech*, 5:255-261, 1987; Tomme et al., *Eur. J. Biochem.*, 170:575-581, 1987; Ong et al., *Tibtech*, 7:239-243, 1989; Svensson et al., *Biochem. J.*, 264:309-311, 1989). The substrate recognition domains are generally linked to the catalytically active core proteins by linker peptides of about 30-40 amino acids.

In their natural environment the function of the hinge is apparently to act as a flexible spacer separating the two functional domains by a distance of over 100 angstroms. Low angle X-ray analysis of intact cellulases indicates that the hinge region has an almost fully extended conformation (Schmuck et al., *Biotechnol. Lett.*, 9:397-402, 1986). This hinge peptide sequence is not included in any of the above mentioned compilations of potential linker regions for the construction of novel fusion proteins.

Rational design of novel fusion proteins with two or more domains will require the availability of relatively short and flexible spacer or linker peptides for joining the two proteins or domains without interfering with their function or with their secretion. This will be particularly important for the design and development of fusion proteins in which the association and interaction of two domains is required for activity such as single chain antibodies and which are designed for secretion from the host cells.

Objects of the Invention

One object of the present invention is to provide a method for producing single chain functional proteins that can be secreted in active form into the periplasmic space of bacteria such as E. coli or into culture media.

Another object of the invention is to produce secretable single chain antibodies.

Another object of the invention is to produce multi-functional secretable single chain proteins.

Another object of the invention is to provide isolated DNA sequences for flexible spacer peptides which are readily adaptable to recombinant methodology and easily used to construct rDNA sequences for single chain proteins whose linked functional domains can interact or assemble to form an active conformation of the protein.

Upon further study of the specification, drawings and appended claims, further objects and advantages of this invention will become apparent to those skilled in the art.

Brief Description of the Drawings

The figures of the constructions are not in scale.

Fig. 1 shows the schematic presentation of the CBH I hinge cloning.

Fig. 2 shows the sequences of the original CBH I hinge cDNA and the synthetic CBH I hinge linker used for the cloning of the single-chain antibodies.

Fig. 3A shows the construction of the plasmid pMLL1.

Fig. 3B shows the construction of the plasmid pMLL2.

Fig. 4 shows the nucleotide and the amino acid sequence of the V_H- linker-V_L insert of the plasmid pML5.

Fig. 5 shows the construction of plasmid pML5 used for expression of the single-chain antibody binding oxazolone with CBH I hinge linker.

Fig. 6 shows the result of the western blot analysis after overnight induction of the pML5 in E. coli RV308 and MC1061 strains. Ox Fab has been used as a positive and pKK223-lacI as a negative control (s=supernatant and c=cells). The difference between the samples s1 and s2 and also between the samples s11 and s22 is only the dilution factor.

Fig. 7 shows the cloning procedure for the CBHII signal sequence.

Fig. 8 shows the construction of plasmid pDS1 containing the CBHII signal sequence preceding the anti-Ox scAb.

Fig. 9 shows the construction of the plasmid pDS2 in which the scAb construction of Fig. 8 is linked to the yeast expression vector pMA91.

Summary of the Invention

The aforementioned objects of the invention and other objects are achieved by providing a method for producing a single chain fusion protein, comprising the steps of:

(a) constructing an rDNA molecule coding for a secretable single chain fusion protein, said fusion protein comprising a plurality of functional proteins or protein domains joined by spacer peptides;

wherein each said spacer peptide comprises a flexible hinge region of a naturally secreted multidomain protein or an analogue thereof;

wherein each adjacent pair of functional proteins or protein domains is joined by a spacer peptide which is not naturally continuous with either functional protein or protein domain to which it is joined; and

wherein at least two of said functional proteins or protein domains each independently exhibit physicochemical or biological activity or interact to form an assembly that exhibits physicochemical or biological activity;

(b) inserting said rDNA molecule into an expression vector capable of transforming a host organism, wherein said vector further comprises DNA sequences coding for a promoter sequence, a signal sequence, a terminator sequence, a replication origin and at least one selection marker for selecting

transformed host cells, and recovering the resultant vector;

(c) transforming a host organism with said resultant vector, and isolating transformants, said host being capable of expressing said single chain fusion protein and secreting it into the periplasmic space of a bacterial host or into the culture medium; and

(d) culturing said transformants in a culture medium, under secretory conditions, and recovering said single chain fusion protein from said periplasmic space or culture medium.

Single chain secretable functional fusion proteins are provided using the foregoing method, as are rDNA molecules, vectors and transformed host organisms suitable for use therein. In particular, single chain antibodies can be produced and secreted in active antigen-binding form using the method of the invention.

Detailed Description

As used herein, the terms, "immunoglobulin", "heavy chain", "light chain" and "Fab" are used in the same way as in *Eur. Pat. Appln. No.* 0125023, of Cabilly et al., published 11/14/84, corresponding to *U.S. Serial No.* 483,457, filed 4/8/83. "Fv" is used to describe those regions of the immunoglobulin molecule which are responsible for antigen or hapten binding. Usually these consist of approximately the first 100 amino acids of the N-termini of the light and the heavy chain of the immunoglobulin molecule.

"Domain" is used to describe an independently folding part of a protein which may or may not function independently. General structural definitions for domain borders in natural proteins are given in Argos, 1988 (Argos, *Protein Engineering*, 2:101-113, 1988).

"Single chain" is used to define a molecule in which two or more proteins or protein domains of the same or different origin are joined together to form a continuous amino acid chain synthesized from a single mRNA molecule (transcript).

"Linker" or "linker peptide" is used to describe an amino acid sequence that extends between adjacent domains in a natural or engineered protein. Linkers useful in the present invention are flexible so that they can be inserted between a variety of proteins or protein domains and used to link these together without interfering with their function. It will be appreciated by the skilled artisan that recombinant DNA techniques can be used to position any such linker peptide sequence between any two domains and that, in particular instances, the linker sequences may be augmented or truncated by a few amino acids in order to facilitate cloning procedures.

"Hinge region" is a term used to describe a class of flexible linker peptide sequences that link functional domains in a protein chain, e.g., the peptide sequences that link the substrate binding domain to the catalytic domain in a glycosidase enzyme such as a cellulase or a glucoamylase.

A general method is presented herein for constructing fusion proteins in which two or more functional proteins or protein domains from the same

or different origin are joined together by a flexible linker peptide to form a single chain fusion protein which can be secreted from bacteria, yeasts and fungi and other eukaryotic hosts. Specifically a method is introduced for constructing single chain antibodies which can be secreted to the culture medium of the above mentioned hosts or to the periplasmic space of bacteria such as E. coli.

Linker peptides from naturally secreted multidomain proteins, e.g., glycosidases such as cellulases, glucoamylases, and other secreted proteins are used as spacers between the proteins or protein domains to be joined. The hinge peptide may be heterologous to the functional proteins or protein domains to which it is joined or homologous to one or both, but will not be the continuation of the natural amino acid sequence for the functional protein/domain into the sequence for the hinge region of the same naturally secreted protein.

Naturally secreted multidomain proteins are understood to embrace any secretable protein having two or more structurally and/or functionally distinct domains. The domains can be biologically active, either individually, e.g., a substrate binding region of an enzyme, or by association of more than one domain, e.g., immunoglobulin (Ig) heavy and light chain variable regions. The activity can be physicochemical, e.g., conferring certain properties upon the protein such as solubility or adsorption onto a solid support or into a membrane or fiber matrix.

The present inventors have now found that these hinge regions are suitable for linking together proteins or protein domains even when close packing of

the two parts is required as in the case of scAbs. Use of the hinge regions of glycosidases or other secreted proteins to link together the hapten binding domains of immunoglobulins, according to the present invention, results in fusion proteins, secretable to the culture medium of various host organisms, and retaining their antigen-binding properties. It will be appreciated that minor variations in the structure of such linker peptides will normally not alter their basic characteristics and destroy their efficacy as linkers that are compatible both with function and secretion of the peptide of which they are a part.

The constructions are made using recombinant-DNA technology and genes and/or cDNAs of the relevant enzymes and other proteins. The DNA sequences coding for the desired fusion protein are inserted in an appropriate expression vector which subsequently transforms the host organism for expression and secretion. The functional portions of the constructs can be selected from among sequences taken from antibodies, enzymes, hormones, viral envelope glycoproteins or the like.

The linkers used in these constructs have as their common characteristic that they facilitate, or at least do not interfere with, secretory processes that transport transcribed proteins to the bacterial periplasmic space or entirely out of eukaryotic cells and into the surrounding medium. Illustrative linkers are the highly conserved regions of glycosidases that contain mainly proline and serine/threonine residues. The hinge regions of most bacterial cellulases and, e.g., IgA, consist of Thr and/or Ser and Pro residues exclusively while in the hinges of fungal cellulases

and glucoamylases, occasional Ala, Cys, Gly, Gln or charged residues are permitted in addition to the Ser, Thr and Pro residues.

Without being bound by any theoretical rationale as to why such linkers are so useful for linking proteins and/or protein domains into single chain secretable proteins, it appears that such linker sequences share the properties of flexibility and substantial absence of secondary structure such as pleated sheet or helical segments. The Thr and Ser residues tend to be O-glycosylated even in the bacterial cellulase hinge regions while Asn residues, which are sites for N-glycosylation, are usually not found in these hinges. This suggests that the O-glycosylation might facilitate the secretion or otherwise be important for the function of these hinges. However, glycosylation of the hinge appears not to be an absolute requirement for the secretion as scAb produced in E. coli is secreted.

Similarly the hinge regions can be used to construct many other types of fusion proteins such as enzyme-enzyme fusions (bifunctional enzymes) in which two or more enzymatic activities are required in a single polypeptide chain or enzyme-binding protein fusions in which the enzyme is bound to a carrier or support by a specialized binding domain in such a way that its biological activity is not destroyed. Such techniques are much wanted in biotechnology to facilitate protein purification or immobilization.

Another example is a fusion protein in which the domains or proteins linked together interact or pack together, and this interaction is required for their correct folding, stability or activity. In

addition to scAbs, an example of such a fusion protein of practical importance is the production of the glycosylated envelope proteins of animal viruses to be used in diagnostics or as vaccines. These envelope glycoproteins contain important antigenic epitopes towards which neutralizing antibodies are elicited (Webster et al., *Nature*, 296:115-121, 1982; Bruck et al., *Virology*, 122:352-362, 1982; Ho-Terry et al., *Arch. Virol.*, 90:145-152, 1986). For instance the two envelope proteins of rubella virus must form a heterodimer in order to be secreted efficiently. For efficient production of these two proteins for medical or diagnostic purposes it would clearly be an advantage if the two proteins involved in the dimerization could be produced as a single polypeptide chain rather than independently. In this way the proportions of the two counterparts would be automatically equal and their secretion and folding would be facilitated by the close presence of the counterpart. The use of the flexible hinge described in this invention also allows the construction of this kind of fusion protein.

It is also possible to make constructs in which more than two proteins or protein domains are linked to a single chain fusion protein by the flexible hinges. For instance a binding domain for cellulose can be linked to either terminus of a desired scAb, thereby facilitating its immobilization on cellulose. In linear form (from N-terminus to C-terminus) such a construct could, e.g., contain: one of the hypervariable domains of the antibody - hinge I - the other hypervariable domain of the antibody - hinge II - the cellulose binding domain. In such a

fusion protein the first hinge allows the packing of the two antibody hypervariable domains onto each other while the second hinge (the same or different from the first one) acts as a spacer to separate the binding domain from the scAb.

Furthermore, it is possible to construct fusion proteins consisting of even more than three domains. For instance bifunctional single chain antibodies could be constructed by linking together two domains of one antibody and two domains of another antibody and subsequently joining these together by yet another hinge. Such a construct thus contains four immunoglobulin domains linked to a single polypeptide via three hinges which allow pairwise packing of the correct immunoglobulin domains and the separation of the two scAbs to allow their independent functioning. Similar multidomain proteins exist in nature, e.g., regulatory proteins and the use of the flexible secretable hinges described in this invention now makes possible the construction of novel multidomain proteins by protein engineering.

The genes coding for the desired proteins or protein domains to be linked are isolated and sequenced using standard recombinant-DNA methods. Alternatively the cDNAs corresponding to the said proteins or protein domains are synthesized enzymatically using the specific mRNAs as templates also using standard recombinant-DNA techniques. The DNA sequences can also be synthesized chemically according to a pre-existing sequence or a sequence designed de novo. The DNA sequences of such proteins or protein domains can be synthesized in such a way as

to optimize the codon usage to correspond to that preferred by the host organism used for expression.

The isolated genes or synthesized cDNAs are linked to suitable expression vectors comprising: 1) an efficient promoter; 2) an efficient signal sequence; and 3) terminator sequences. In E. coli, many powerful promoters have been described, e.g., lac, trp or phage lambda Pr and Pl promoters as well as hybrid promoters such as tac. In yeast, powerful promoters such as the phosphoglycerokinase (PGK) or alcohol dehydrogenase (ADH) promoters have been used, and in filamentous fungi promoters of, e.g., hydrolytic enzymes such as the cellulase, cellobiohydrolase I, of Trichoderma reesei have been used. Any strong promoter functional in the host organism used can in principle be used for the expression of the secretable fusion proteins. It is an advantage, although not an absolute requirement, for the promoter to be kept in a repressed state until it is activated (induced) at a given time point by, e.g., elevated temperature, as in the case of the lambda Pr and Pl promoters, or by adding particular inducing agents such as IPTG in the case of the lac or tac promoters of E. coli.

The signal sequences directing secretion to the periplasmic space or to the culture medium of E. coli include those of OmpA protein of E. coli or pectate lyase of Erwinia carotovora but a number of other signal sequences also can be used. In yeast, secretion of foreign proteins can be achieved using the signal sequences of the yeast alpha factor, acid phosphatase or invertase or other signal sequences of the rare secreted proteins of yeast. Also, signal sequences from fungal cellulases have worked in yeast

(Penttilä et al., *Yeast*, 3:75-83, 1987; Penttilä et al., *Gene*, 63:103-112, 1988) and other eukaryotic and possibly prokaryotic signal sequences may be used in yeast as well. In filamentous fungi, the signal sequences of its cellulases have been shown to work with other proteins and signal sequences from other secreted fungal proteins can also be used. Thus, for any selected organism, a signal sequence can be chosen from among those derived from homologous or heterologous secreted proteins.

The terminator sequences are isolated from genes homologous to each production organism. In E. coli, the terminators of, e.g., rrnB from the ribosomal RNA transcription terminators work well. In yeast, terminators of yeast proteins or, e.g., terminators of filamentous fungal genes can be used.

The 5' end of the gene or cDNA of the first protein or domain is cloned in frame with the 3' end of the signal sequence chosen to generate a proper signal sequence cleavage site. This may require alteration of the DNA and/or amino acid sequence of the first protein. The DNA corresponding to the secretable joining peptide is isolated from a gene or cDNA coding for a secretable protein or synthesized and joined at its 5' end to the 3' end of the DNA of the first protein. Finally, the DNA for the second protein or domain is cloned in between the 3' end of the linker DNA and the 5' end of the terminator.

The plasmids used contain a replication origin functional in the production organism and selection markers for the selection of transformed cells from nontransformants.

Examples of usable host organisms include bacteria, e.g., Escherichia coli MC1061, derivatives of Bacillus subtilis BRB1 (Sibakov et al., *Eur. J. Biochem.*, 145:567-572, 1984), Staphylococcus aureus SA1123 (Iordanescu, *J. Bacteriol.* 12:597-601, 1975) or Streptococcus lividans (Hopwood et al., Genetic Manipulation of Streptomyces, a Laboratory Manual, The John Innes Foundation, Norwich 1985); yeasts, e.g., Saccharomyces cerevisiae AH 22 (Mellor et al., Gene, 24:1-14, 1983) and Schizosaccharomyces pombe; filamentous fungi, e.g., Aspergillus nidulans, Aspergillus awamori (Ward, *Proc. Embo-Alko Workshop on Molecular Biology of Filamentous Fungi*, Helsinki, pp. 119-128, 1989), Trichoderma reesei (Penttilä et al., *Gene* 61:155-164, 1987; Harkki et al., *Bio/Technology*, 7:596-603, 1989), the foregoing being illustrative but not limitative of the many possible host organisms know to the art. In principle, all hosts capable of secretion can be used whether prokaryotic or eukaryotic.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever. In the following examples, all temperatures are set forth uncorrected in degrees Celsius; unless otherwise indicated, all parts and percentages are by weight.

Example 1

ScFv in E. coli

The following plasmids are used to make the construct:

Plasmids

pLT1 (see Fig. 4)
pML2 (see Fig. 1)
pML3 (see Fig. 7)
pML4 (see Fig. 7)
pML5 (see Fig. 7)
pMLL1 (see Fig. 3)
pMLL2 (see Fig. 4)
pTI8 (see Fig. 5)
pSP72 (Promega)
pSP73 (Promega)

E. coli strains DH5α (F-, endA1, hsdR17(r_k-, m_k), supE44, thi-1,λ-, recA1, gyrA96, relA1, Δ(argF-laczya)U169, φ80dlacZΔM15), MC1061 (F-,araD139, Δ(ara-leu)7696, ΔlacI74, galU-, galK-, hsr-, hsm+, strA) and RV308 (su-,ΔlacX74, galSII::OP308, strA) are used as hosts in transformation and expression studies.

Table I shows the primers annealing to the sense strand in their complementary form, the relevant amino acid residues incorporated to the primers and the restriction sites (bolded) used for the in-frame fusions. The restriction sites used for the subcloning of the amplified fragments are also indicated.

Table II shows the production levels of the active single-chain antibody measured by ELISA.

Cloning of the CBH I Hinge Linker

DNA manipulations were mainly performed according to standard protocols (Maniatis et al., Molecular cloning. A laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982). The cloning procedure for the CBH I hinge linker is schematically presented in Figure 1. Oligomers (45- and 81-mers) were synthesized on an Applied Biosystems DNA synthesizer (Model 391 PCR MATE) using the phosphoramidite chemistry. The nucleotide sequence of the authentic CBHI hinge-region is shown in Fig. 2 as well as the nucleotide sequence of the synthetic CBHI hinge-linker. The synthesized oligonucleotides were provided with restriction enzyme cleavage sites used in the cloning steps. The partial overlapping oligomers were denatured, annealed and extended using Klenow polymerase as follows. The hinge oligomers, 100 pmol of the 45-mer and 100 pmol of the 81-mer, were denatured in 5 µl of water at 65°C for 5 min. The oligomers were annealed in 10 µl of 30 mM Tris-HCl (pH 7.5), 150 mM NaCl, 30 mM $MgCl_2$, 15 mM DTT and 0.1 mg/ml gelatin for 2 h at 37°C and then allowed to cool slowly to room temperature. For the extension at 37°C for 2 h the annealed oligo mixture was diluted to a final volume of 30 µl containing 300 µM deoxynucleoside triphosphates (dNTPs) and 5 U of Klenow fragment of E. coli DNA polymerase I (Boehringer). The reaction mixture was extracted with phenol, ammonium acetate was added into the water phase to a final concentration of 2 M before ethanol precipitation. The double-stranded hinge fragment was cleaved with 30 units of XbaI and SacI at 37°C

overnight, phenol extracted, ethanol precipitated, lyophilized and dissolved into 10 µl of water.

The double-stranded hinge fragment was cloned into a pSP72 vector, from which the XhoI restriction site was deleted (Fig. 1). The hinge linker contained an internal XhoI restriction site, which was used for the in-frame fusion of the V_L fragment. Therefore the XhoI restriction site was deleted by cutting the pSP72 vector with NdeI and XhoI and the 5'-protruding ends were filled in with 2 U of Klenow fragment and 40 µM dNTPs at room temperature for 30 min. The reaction mixture was extracted with phenol and precipitated with ethanol. The deleted linear pSP72ΔXhoI vector was separated from the short XhoI-NdeI fragment (87 bp) on 1% agarose gel and eluted using GeneClean kit (BIO 101). The linear vector was recircularized with 1 U of T4 DNA ligase (Boehringer) in 10 µl containing 50 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 5 mM DDT, 100 µM spermidine, 100 µM EDTA and 10 mM ATP at 15°C overnight. An aliquot of 1 µl was transformed into E. coli DH5α strain according to Hanahan (Hanahan, *J. Mol. Biol.*, 166:557-580, 1983).

Plasmid pSP72ΔXhoI was cleaved with SacI and XbaI and ligated with 1 µl and 1 µl of 1:20 dilution of the double-stranded SacI-XbaI hinge fragment. The authenticity of the cloned hinge insert was confirmed by sequencing using an oligonucleotide primer complementary to the T7 promoter of the vector using sequenase (USB) according to manufactures instructions. The hinge construction was designated pML2.

Cloning of the Ox V_H and V_L Domains

For the in-frame fusions with the hinge coding region the 5' end of the Ox V_H cDNA and the 3' end of the V_L cDNA were modified by using the polymerase chain reaction (PCR) method (Saiki et al., *Science*, 239:487-494, 1988; Table I). The V_H region was amplified using pTI8 construction and the V_L region using pLT1 construction as templates. Plasmid pTI8 contains the cDNA fragments of the Fab coding regions of the heavy and light chains of an anti-2-phenyloxazolone binding antibody, IgG1 subclass (OX-1 IgG1). Plasmid pLT1 contains the cDNA of the complete light (κ) chain coding region cloned into pSP73 vector. The oligonucleotide primers used for the amplifications of single-chain antibody are shown in Table I. Recognition sequences for EcoRI and HindIII were incorporated into the primers to allow the subcloning of the amplified fragments into EcoRI-HindIII cleaved pSP72 and pSP73 vectors (Fig. 3a and Fig. 3b). The amplification was done in 100 μl volume containing 50 ng of template DNA, 100 pmol of each primer, 200 μl dNTPs, 10 mM Tris-HCl (pH 8.0), 10 mM $MgCl_2$, 50 mM KCl, 1 mg/ml gelatin, 1 mM ß-ME and 4 U of Taq I polymerase (USB). The reaction mixtures were amplified with 25 repeating cycles of denaturation at 94°C for 1 min., annealing at 55°C for 1 min. and polymerization at 72°C for 3 min. The reaction mixture was extracted with phenol and precipitated with ethanol. The amplified V_H and V_L fragments were cleaved with EcoRI and HindIII and purified and ligated into pSP73 and pSP72 vectors. These constructions are designated pMLL1 and pMLL2. The nucleotide sequences of the amplified DNA fragments were confirmed by sequencing

using primers complementary to both T7 and SP6 promoters (Fig. 4).

Construction of the Ox V_H-Hinge-V_L E. coli Expression Vector

The final cloning procedure for the Ox scFv (pML5) is presented in Figure 5. The pML2 construction, containing the CBH I hinge insert, was cleaved with XhoI and HindIII and ligated with the XhoI-HindIII V_L fragment derived from the pMLL2 construction (Fig. 3b). The V_H fragment was cleaved from the pMLL1 construction (Fig. 3a) with EcoRI and XmaI and ligated with the pML3 cleaved with the same restriction enzymes. The resulting construction, designed pML4, contains the signal sequence of Erwinia carotovora's pectate lyase (pelB) gene (Lei et al., J. Bact., 169: 4379-4383, 1987; Better et al., Science, 240: 1041-1043, 1988) and the whole coding region of the V_H-hinge-V_L fusion. The EcoRI-HindIII insert of pML4 was ligated into the E. coli pKK233-3 expression vector (Pharmacia) under the control of the tac promoter (Fig. 5). To achieve efficient repression of the tac promoter before induction with IPTG the $lacI^q$ gene fragment was cloned into the expression vector (Brosius, Gene, 27:161-172, 1984; Stark, Gene, 51:255-267, 1987; Amann et al., Gene 69: 301-315, 1988). The OX-1 single-chain antibody expression construction obtained by this procedure is designated pML5.

Expression of Single-Chain Antibody Genes

Plasmid DNA from several clones of pML5 was transformed into E. coli strains RV308 and MC1061. For expression studies transformant colonies were

inoculated into LB medium containing 100 µg/ml ampicillin and the cultures were grown at 37°C overnight. After 1:50 dilution of a 50 ml culture was incubated at 30°C with shaking 115 rpm until the A_{600} was between 0.5 and 1.5 before the induction with the 1 mM isopropyl ß-D-thiogalactopyranoside (IPTG). As a negative control the expression vector without the insert was induced identically. The growth of the induced culture was followed by measuring the cell densities at different time points after the induction.

For the activity analysis of the OX-1 single-chain antibody, aliquots of 2 ml of the culture were centrifuged. The culture media were concentrated approximately 20 times using Centricon 10 (Amicon) concentrator units and an equal volume of 2x Laemmli's (1970) sample buffer was added. The cell pellets were resuspended into the same sample buffer and the samples were boiled for 3 min and 15 µl of them were subjected to 15% SDS-PAGE analysis according to Laemmli (1970).

The proteins were transferred electrophoretically onto nitrocellulose filters (Towbin et al., *Proc. Natl. Acad. Sci. USA*, 76:4350-4354, 1979) and the single-chain antibody was detected with a polyclonal antiserum raised in rabbits against the Fab-moiety of the oxazolone monoclonal antibody (Ox IgG3 Fab antiserum, Fig. 6). The ELISA assay was essentially performed as described by Hudson and Hay (1976). Oxazolone coupled to BSA in a ration of 22:1 (Ox-BSA) was used in the assay to detect immunoreactive single-chain antibody. The antigen was adsorbed to 96-well microtiter plates (NUNC Immunoplate I) by incubating

100 µl/well of a Ox-BSA solution containing 46 µg/ml in 0.1 M carbonate buffer (pH 9.6) at 4°C overnight. Unless indicated otherwise in the text, all the following binding steps were carried out by applying 100 µl of the respective solution per well and shaking the plates (Wallac Plateshake) at room temperature for 1 h. After each binding step the plates were washed three times in a Platewash (Wallac) with PBS containing 50 mM phosphate buffer (pH 7.4) and 0.9% NaCl. Free adsorption sites in the wells were blocked by incubating 250 µl/well of 0.5% BSA in PBS (BSA-PBS). Supernatant samples were taken from cultures of E. coli RV308/pML5 and MC1061/pML5 after different times of induction and appropriate dilutions in BSA-PBS were dispersed in duplicate into the antigen-coated wells and incubated for 2 h. Bound single-chain antibody was detected by subsequent incubations with first Ox IgG3 Fab antiserum in BSA-PBS and then with alkaline phosphatase-conjugated goat anti-rabbit IgG (H+L) (BioRad) in BSA-PBS. After addition of substrate solution containing 2 mg/ml p-nitrophenylphosphate in diethanolamine-$MgCl_2$-buffer (Orion) the absorbance in each well was read using a Multiskan MCC/340 (Labsystems) with an interference filter of 405 nm. The amount of immunoreactive single-chain antibody in each well was determined by comparing the absorbances derived from the samples to a standard curve obtained by incubating known concentrations of the purified isotype monoclonal antibody on the same microtiter plate (Table II).

TABLE I

VH 1 : NH_2-terminal primer of VH, COMPLEMENTARY TO THE ptac PROMOTER REGION

5' GCGCCGACATCATAACGGTCC 3'

VH 2 : COOH-terminal primer of VH

5' ACTTGAAGCTTCCCGGGTGCAGAGACAGTGACCAGAGTCCC 3'

3' TGAACTTCGAAGGGCCCACGTCTCTGTCACTGGTCTCAGGG 5'

VH ⟷ Hinge

G T L V T V S A | P G

5' GGGACTCTGGTCACTGTCTCTGCACCCGGGAAGCTTCAAGT 3'

XmaI

VL 1 : NH_2-TERMINAL PRIMER OF VL

Hinge ⟷ VL

S S P G P T Q S H Y | Q I V

5' GATTAGAATTCCTCGAGTCCAGGGCCCACCCAGTCTCATTATCAAATTGTT

XhoI

L T Q S P

CTCACCCAGTCTCCAG 3'

VL 2 : COOH-TERMINAL PRIMER OF VL

5' ACATCAAGCTTCTATTTCAGCTCCAGCTTG 3'

3' TGTAGTTCGAAGATAAAGTCGAGGTCGAAC 5'

VL ↰

K L E L K * |

5' CAAGCTGGAGCTGAAATAGAAGCTTGATGT 3'

HindIII

TABLE II

Production levels of active anti Ox ScFv in the culture medium of E. coli as measured by ELISA

plasmid/ strain	anti Ox ScFv (mg/l)
pML5/ RV308	3.4
pML5/ MC1061	0.3
pKK223-lacI9/ RV308	0

Example 2

Anti Ox ScAb in yeast

The following plasmids were used to make the construct:

Plasmids

pTTc9 (Teeri et al., *Gene*, 51:43-52, 1987)
pMA91 (Mellor et al., *Gene*, 24:1-14, 1983)
pMLL4 (see Example 1)
pSP72 (Promega)

The Saccharomyces cerevisiae strain AH 22 (leu2-3 leu2-112 his4-519 can1 gal2 cin+)(Mellor et al., 1983, *supra*) was used as host for transformation and expression studies.

Yeast transformation was carried out according to Keszenman-Pereyra et al., *Curr. Genet.*, 13:21-23, 1988) and the transformants were grown in synthetic complete medium lacking leucine (Sherman et al., *Methods in Yeast Genetics, A Laboratory Manual*. Cold Spring Harbor, USA, 1983)

Cloning of the Signal Sequence

DNA manipulations are performed essentially as described in Maniatis (Maniatis et al., 1982, *supra*) and Innis (Innis et al., *PCR Protocols. A Guide for Methods and Applications*, Academic Press, San Diego, CA., 1990). The cloning procedure of the CBHII signal sequence is represented schematically in Figure 7. The DNA sequence for the first oligonucleotide primer for fragment amplification is derived from the region upstream of the EcoRI site of the cbh2 cDNA in pTTc9. The second oligonucleotide primer contains the DNA

sequence coding for the last four amino acids at the CBHII signal sequence and the first four amino acids of the anti Ox L-chain as well as the HindIII restriction enzyme recognition sequence as shown in Figure 7. The fragment is amplified using the protocols described in Example 1 and subsequently digested with the restriction enzymes EcoRI and HindIII. The digested fragment is cloned into the plasmid pSP72 digested with EcoRI and HindIII to generate the plasmid pDS1 (Figure 8).

Construction of the Yeast Expression Vector

The DNA fragment corresponding to the anti-Ox scAb preceded by the cBHII signal sequence is released from pDS1 by cleavage with EcoRI and HindIII, the protruding ends are blunted using S1 nuclease, and the fragment is cloned in the yeast expression vector pMA9I cleaved with BglII and treated with S1 nuclease to remove the protruding ends. The resulting plasmid pDS2 (Fig. 9) is transformed to E. coli DH5α. The plasmid DNA is isolated and the construction is checked by sequencing.

Expression of the ScAbs in yeast

The plasmid pDS2 is transformed into yeast. The transformants are grown as described by Penttilä et al. (Penttilä et al., *Yeast*, 3:75-83, 1987). After the growth the yeast cells are removed by centrifugation and the activity of the anti-Ox scAbs assayed as in Example 1. Functional single chain antibody is found in the culture medium.

From the foregoing description, one of ordinary skill in the art can easily ascertain the

essential characteristics of this invention and, without departing from the spirit and scope thereof, can make various changes and modifications of the invention to adapt it to various usages and conditions.

WHAT IS CLAIMED IS:

1. A recombinant DNA (rDNA) molecule coding for a secretable single chain fusion protein, said fusion protein comprising a plurality of functional proteins or protein domains joined by spacer peptides;

wherein each said spacer peptide comprises a flexible hinge region of a naturally secreted multidomain protein or an analogue thereof;

wherein each adjacent pair of functional proteins or protein domains is joined by a spacer peptide which is not naturally continuous with either functional protein or protein domain to which it is joined; and

wherein at least two of said functional proteins or protein domains each independently exhibit a physicochemical or biological activity or interact to form an assembly that exhibits a physicochemical or biological activity.

2. The rDNA molecule of claim 1, coding for a secretable single chain fusion protein comprising two functional proteins or protein domains joined by a single spacer peptide.

3. The rDNA molecule of claim 2, wherein said fusion protein is a single chain antibody.

4. The rDNA molecule of claim 3, wherein the functional domains of said fusion protein comprise the antigen-binding regions of the heavy and light chains of an antibody or antibody fragment.

5. The rDNA molecule of claim 1, wherein one or more of the peptide spacers of said fusion protein is the hinge region of a glycosidase enzyme.

6. The rDNA molecule of claim 5, wherein said glycosidase is a cellulase or a glucoamylase.

7. The rDNA molecule of claim 6, wherein said glycosidase is cellobiohydrolase I.

8. The rDNA molecule of claim 1, wherein one or more of the peptide spacers of said fusion protein is the hinge region of an immunoglobulin.

9. The rDNA molecule of claim 1, wherein each of said peptide spacers of said fusion protein is heterologous to each of the proteins or protein domains to which it is joined.

10. The rDNA molecule of claim 1, wherein the functional proteins or protein domains of said fusion protein are selected from the group consisting of antibody, enzyme, hormone and viral envelope glycoprotein functional proteins or protein domains.

11. An isolated single chain fusion protein coded for by the rDNA molecule of any of claims 1-10.

12. An elongated rDNA molecule, comprising the rDNA sequence of the rDNA molecule of claim 1 coding for said single chain fusion protein, and further comprising one or more sequences selected from the group consisting of a promoter sequence, a signal sequence and a terminator sequence.

13. A recombinant expression vector capable of transforming a host organism, wherein said vector comprises the elongated rDNA molecule of claim 12, including said rDNA sequence coding for said single chain fusion protein and DNA sequences coding for a promoter sequence, a signal sequence and a terminator sequence, and wherein said vector further comprises a replication origin and at least one selection marker for selecting transformed host cells.

14. A host transformant transformed by the expression vector of claim 13 and capable of expressing and secreting said single chain fusion protein.

15. The transformed host of claim 14 which is selected from the group consisting of bacteria, yeast and filamentous fungi.

16. The transformed host of claim 14 which is a eukaryote.

17. The transformed host of claim 14, wherein said host is selected from the group consisting of Escherichia coli, Bacillus subtilis, Staphylococcus aureus, Streptococcus lividans, Saccharomyces cerevisiae, Schizosaccharomyces pombe, Aspergillus nidulans, Aspergillus awamori and Trichoderma reesei.

18. The transformed host of claim 17, wherein said host is E.coli transformed by an expression vector comprising a promoter selected from the group consisting of lac, trp, phage lambda Pr, phage lambda Pl and tac promoters.

19. The transformed host of claim 17, wherein said host is E. coli transformed by an expression vector comprising the signal sequence for OmpA protein or the signal sequence for the pectate lyase of Erwinia carotovora.

20. The transformed host of claim 15, wherein said host is yeast transformed by an expression vector comprising the PGK or ADH promoter.

21. The transformed host of claim 15, wherein said host is yeast transformed by an expression vector comprising the yeast alpha factor, acid phosphatase or invertase signal sequence.

22. The transformed host of claim 15, wherein said host is yeast transformed by an expression vector comprising the fungal cellulase signal sequence.

23. The transformed host of claim 15, wherein said host is a filamentous fungus transformed by an expression vector comprising a cellulase promoter.

24. The transformed host of claim 23, wherein said cellulase is the cellobiohydrolase I of Trichoderma reesei.

25. The transformed host of claim 15, wherein said host is a filamentous fungus transformed by an expression vector comprising a cellulase signal sequence.

26. A method for producing a single chain fusion protein, comprising the steps of:

(a) constructing an rDNA molecule coding for a secretable single chain fusion protein, said fusion protein comprising a plurality of functional proteins or protein domains joined by spacer peptides;

wherein each said spacer peptide comprises a flexible hinge region of a naturally secreted multidomain protein or an analogue thereof;

wherein each adjacent pair of functional proteins or protein domains is joined by a spacer peptide which is not naturally continuous with either functional protein or protein domain to which it is joined; and

wherein at least two of said functional proteins or protein domains each independently exhibit physicochemical or biological activity or interact to form an

assembly that exhibits physicochemical or biological activity;

(b) inserting said rDNA molecule into an expression vector capable of transforming a host organism, wherein said vector further comprises DNA sequences coding for a promoter sequence, a signal sequence, a terminator sequence, a replication origin and at least one selection marker for selecting transformed host cells, and recovering the resultant vector;

(c) transforming a host organism with said resultant vector, and isolating transformants, said host being capable of expressing said single chain fusion protein and secreting it into the periplasmic space of a bacterial host or into the culture medium; and

(d) culturing said transformants in a culture medium, under secretory conditions, and recovering said single chain fusion protein from said periplasmic space or culture medium.

27. The method of claim 26, wherein in step (a)(1), said rDNA molecule codes for a fusion protein which is a single chain antibody.

28. The method of claim 27, wherein the functional domains of said fusion protein comprise the antigen-binding regions of the heavy and light chains of an antibody or antibody fragment.

29. The method of claim 26, wherein in step (a), said rDNA molecule codes for a fusion protein, a peptide spacer of which is the hinge region of a glycosidase enzyme.

30. The method claim 29, wherein said glycosidase is a cellulase or a glucoamylase.

31. The method of claim 30, wherein said glycosidase is cellobiohydrolase I.

32. The method of claim 26, wherein in step (a), said rDNA molecule codes for a fusion protein, a peptide spacer of which is the hinge region of an immunoglobulin.

33. The method of claim 26, wherein in step (a), said rDNA molecule codes for a fusion protein, the functional proteins or protein domains of which are selected from the group consisting of antibody, enzyme, hormone and viral envelope glycoprotein functional proteins or protein domains.

34. The method of claim 26, wherein in step (c), said transformed host is selected from the group consisting of bacteria, yeast and filamentous fungi.

35. The method of claim 26, wherein in step (c), said transformed host is a eukaryote.

36. The method of claim 26, wherein said host is selected from the group consisting of Escherichia coli, Bacillus subtilis, Staphylococcus aureus, Streptococcus lividans, Saccharomyces cerevisiae, Schizosaccharomyces pombe, Aspergillus nidulans, Aspergillus awamori and Trichoderma reesei.

37. The method of claim 36, wherein said host is E. coli, and wherein in step (b), said expression vector comprises a promoter selected from the group consisting of lac, trp, phage lambda Pr, phage lambda Pl and tac promoters.

38. The method of claim 36, wherein said host is E. coli, and wherein in step (b), said expression vector comprises the signal sequence for OmpA protein or the signal sequence for the pectate lyase of Erwinia carotovora.

39. The method of claim 34, wherein in step (c), said host is yeast transformed in step (b) by an expression vector comprising the PGK or ADH promoter.

40. The method of claim 34, wherein in step (c), said host is yeast transformed in step (b) by an expression vector comprising the yeast alpha factor, acid phosphatase or invertase signal sequence.

41. The method of claim 34, wherein in step (c), said host is yeast transformed in step (b) by an expression vector comprising the fungal cellulase signal sequence.

42. The method of claim 34, wherein in step (c), said host is a filamentous fungus transformed in step (b) by an expression vector comprising a cellulase promoter.

43. The method of claim 42, wherein said cellulase is the cellobiohydrolase I of Trichoderma reesei.

44. The method of claim 34, wherein in step (c), said host is a filamentous fungus transformed in step (b) by an expression vector comprising a cellulase signal sequence.

45. An isolated single chain fusion protein produced by the method of any of claims 26-44.

46. An isolated, purified DNA fragment coding for a hinge region peptide of a glycosidase, and optionally further comprising short flanking sequences adapted for use in inserting said fragment into a recombinant DNA construct.

47. The DNA fragment of claim 46, wherein said hinge region peptide is the hinge region of a cellulase or glucoamylase.

48. The DNA fragment of claim 46, wherein said hinge region peptide is the hinge region of an immunoglobulin.

49. A recombinant DNA cloning vector, comprising a promoter, a signal sequence, one or more sequences each coding for a hinge region peptide of a glycosidase flanked at either end by unique multicloning sequences, a transcription terminator, an origin for replication and a selectable marker.

50. The vector of claim 49, wherein said hinge region peptide is the hinge region of a cellulase or glucoamylase.

51. The vector of claim 49, wherein said hinge region peptide is the hinge region of an immunoglobulin.

FIG. 1

ORIGINAL CBH I - HINGE - cDNA:

```
                P  G  G  N  R  G  T  T  T  T  R  R  P  A  T  T  T  G  S  S  P  G  P  T
                CCC GGC GGA AAC CGT GGC ACC ACC ACC ACC CGC CGC CCA GCC ACT ACC ACT GGA AGC TCT CCC GGA CCT ACC
 Q  S  H  Y
CAG TCT CAC TAC
```

SYNTHETIC CBH I - HINGE - LINKER:

```
                P  G  G  N  R  G  T  T  T  T  R  R  P  A  T  T  T  G  S  S  P  G  P  T
TCATAGAGCTC CCC GGG GGT AAC CGC GGT ACC ACC ACT ACC CGT CGT CCG GCT ACC ACC ACT GGC TCG AGT CCA GGG CCC ACC
     Sac I  Xma I                                                                 Xho I
 Q  S  H  Y
CAG TCT CAT TAT TCTAGAAACAT
                Xba I
```

FIG. 2

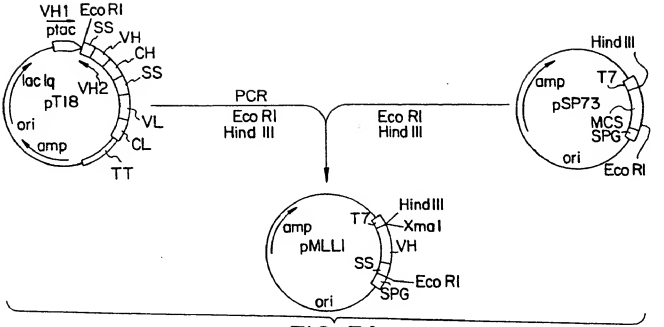

VH1
Eco RI
ptac
SS
VH
CH
SS
lac Iq
VH2
pT18
ori
VL
amp
CL
TT
PCR
Eco RI
Hind III
Eco RI
Hind III
Hind III
amp
T7
pSP73
MCS
SPG
ori
Eco RI
Hind III
amp
T7
Xma I
pMLL1
VH
SS
Eco RI
SPG
ori

FIG. 3A

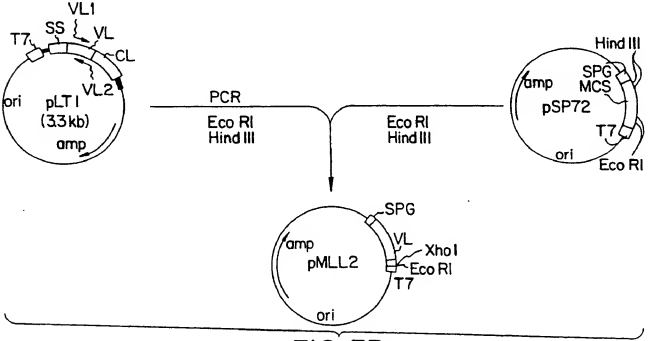

VL1
T7
SS
VL
CL
VL2
ori
pLT1
(3.3kb)
amp
PCR
Eco RI
Hind III
Eco RI
Hind III
Hind III
SPG
amp
MCS
pSP72
T7
ori
Eco RI
SPG
amp
VL
Xho I
pMLL2
Eco RI
T7
ori

FIG. 3B

```
CGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAACAGAATTC
          -50        -40        -30        -20        -10
            SS
 M  K  Y  L  L  P  T  A  A  A  G  L  L  L  L  A  A  Q  P  A
ATGAAATACCTATTGCCTACGGCAGCCGCTGGATTGTTATTACTCGCTGCCCAACCAGCG
1          10         20         30         40         50         60
            VH
 M  A  Q  V  Q  L  K  E  S  G  P  G  L  V  A  P  S  Q  S  L
ATGGCCCAGGTGCAGCTGAAGGAGTCAGGACCTGGCCTGGTGGCGCCCTCACAGAGCCTG
          70         80         90        100        110        120

 S  I  T  C  T  V  S  G  F  S  L  T  S  Y  G  V  Q  W  V  R
TCCATCACTTGCACTGCTTCTGGGTTTTCATTAACCAGCTATGGTGTACACTGGGTTCGC
          130        140        150        160        170        180

 Q  P  P  G  K  G  L  E  W  L  G  V  I  W  A  G  G  S  T  N
CAGCCTCCAGGAAAGGGTCTGGAGTGGCTGGGAGTAATATGGGCTGGTGGAAGCACAAAT
          190        200        210        220        230        240

 Y  N  S  A  L  M  S  R  L  S  I  S  K  D  N  S  K  S  Q  V
TATAATTCGGCTCTCATGTCCAGACTGAGCATCAGCAAAGACAACTCCAAGAGCCAAGTT
          250        260        270        280        290        300

 F  L  K  M  N  S  L  Q  T  D  D  T  A  M  Y  Y  C  A  R  D
TTCTTAAAAATGAACAGTCTGCAAACTGATGACACAGCCATGTACTACTGTGCCAGAGAT
          310        320        330        340        350        360
                                                    CBHI hinge
 R  G  A  Y  W  G  Q  G  T  L  V  T  V  S  A  P  G  G  N  R
CGGGGGGCTTACTGGGGCCAAGGGACTCTGGTCACTGTCTCTGCACCCGGGGGTAACCGC
          370        380        390        400        410        420

 G  T  T  T  T  R  R  P  A  T  T  T  G  S  S  P  G  P  T  Q
GGTACCACCACTACCCGTCGTCCGGCTACCACCACTGGCTCGAGTCCAGGGCCCACCCAG
          430        440        450        460        470        480
            VL
 S  H  Y  Q  I  V  L  T  Q  S  P  A  I  M  S  A  S  P  G  E
TCCCATTATCAAATTGTTCTCACCCAGTCTCCAGCAATCATGTCTGCATCTCCAGGGGAG
          490        500        510        520        530        540

 K  V  T  M  T  C  S  A  S  S  S  V  S  Y  M  H  W  Y  Q  Q
AAGGTCACCATGACCTGCAGTGCCAGCTCAAGTGTAAGTTACATGCACTGGTACCAGCAG
          550        560        570        580        590        600

 K  S  G  T  S  P  K  R  W  I  Y  D  T  S  K  L  A  S  A  V
AAGTCAGGCACCTCCCCCAAAAGATGGATTTATGACACATCCAAACTGGCTTCTGCAGTC
          610        620        630        640        650        660

 P  A  R  F  S  G  S  G  S  G  T  S  Y  S  L  T  I  S  S  M
CCTGCTCTCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTCTCACAATCAGCAGCATG
          670        680        690        700        710        720

 E  A  E  D  A  A  T  Y  Y  C  Q  Q  W  S  S  N  P  L  T  F
GAGGCTGAAGATGCTGCCACTTATTACTGCCAGCAGTGGAGTAGTAACCCACTCACGTTC
          730        740        750        760        770        780

 G  A  G  T  K  L  E  L  K  *
GGTGCTGGGACCAAGCTGGAGCTGAAATAG
          790        800        810
```

FIG. 4

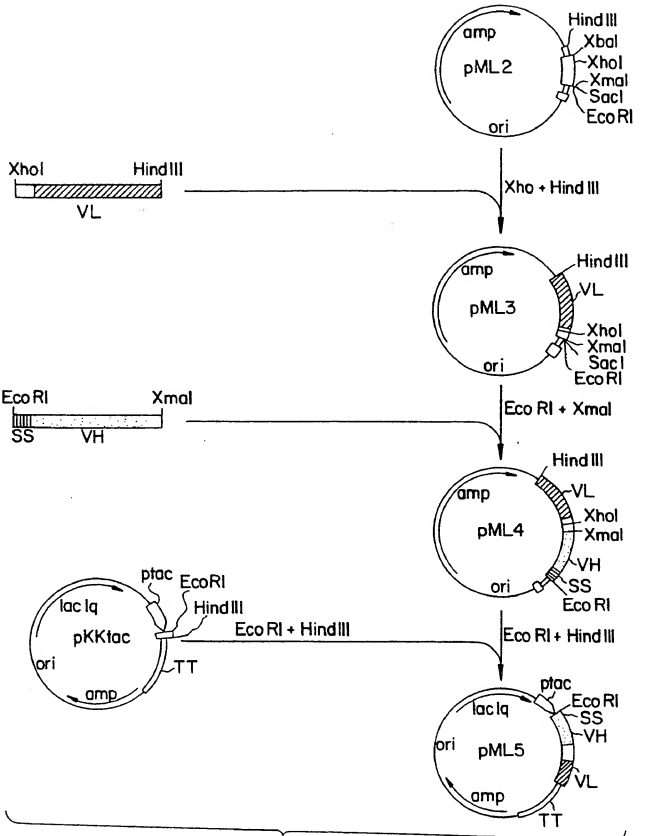


FIG. 5

SUBSTITUTE SHEET

pML5 MC1061/c
pML5 MC1061/s
pML5 RV308/c
pML5 RV308/s
pKK223-1acI(RV308/c
pKK223-1acI(RV308/s
OX-1 Fab
94.0 kD
67.0
43.0
30.0
20.1
RV308
pML5/c
pML5/s2
pML5/s1
pKK223-1acI/c
pKK223-1acI/s
OX-1 Fab
43.0 kD
30.0
20.1
MC1060
pML5/c
pML5/s22
pML5/s11
pKK223-1acI/c
pKK223-1acI/s
OX-1 Fab
43.0 kD
30.0
20.1

FIGURE 6

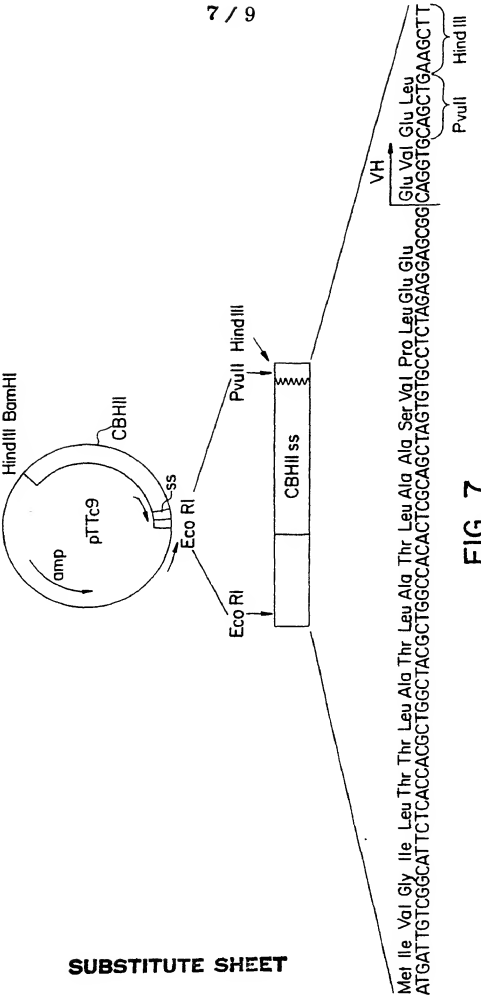


FIG. 7

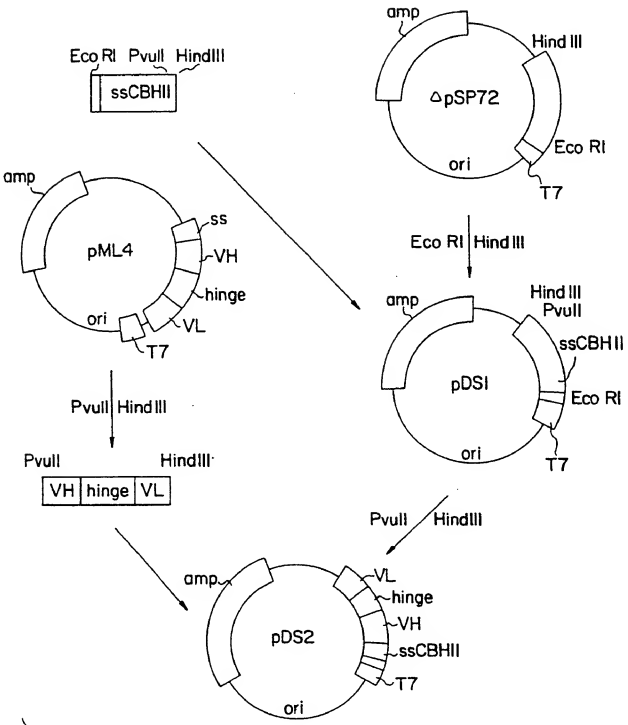


FIG. 8

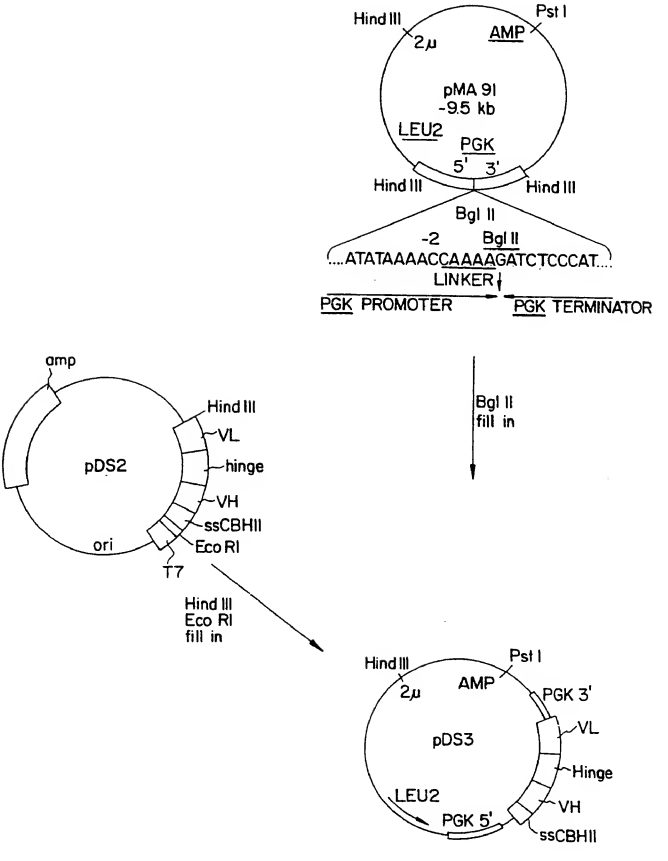


FIG. 9

INTERNATIONAL SEARCH REPORT

International Application No [redacted]/FI 91/00225

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all)

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC5: C 12 N 15/62, 15/13, C 07 K 13/00 // A 61 K 39/395

II. FIELDS SEARCHED

Minimum Documentation Searched[7]	
Classification System	Classification Symbols
IPC5	C 12 N; C 07 K; A 61 K

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in Fields Searched[8]

SE,DK,FI,NO classes as above

III. DOCUMENTS CONSIDERED TO BE RELEVANT[9]

Category *	Citation of Document,[11] with indication, where appropriate, of the relevant passages[12]	Relevant to Claim No.[13]
P,X	Dialog Information Services, File 55, Biosis 85-92, Dialog accession no. 9031001, BIOSIS accession no. 93016001, Takkinen K et al: "An active single-chain antibody containing a cellulase linker domain is secreted by escherichia-coli", Protein Eng 4 (7).1991. 837-842 --	1-11
O,X	KEYSTONE SYMPOSIA ON MOLECULAR & CELLULAR BIOLOGY 20 1991 April 8-14 1991 (Journal of cellular biochemistry Supplement 0 (15 part G) 1991, Tuula T. Teeri et al: "Strategies for secretion of antibody fragments and single chain antibodies in escherichia coli", p 217 Abstract R 344 --	1-11

* Special categories of cited documents:[10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
5th March 1992	1992 -03- 10
International Searching Authority	Signature of Authorized Officer
SWEDISH PATENT OFFICE	Mikael G:son Bergstrand

Form PCT/ISA/210 (second sheet) (January 1985)

III. DOCUMENTS C NSIDERED TO BE RELEVANT

Category *	Citation of Document, with indication, where appropriate, of the relevant passages	Relevant to Claim No
X	GENE, Vol. 93, 1990 Petra Markmeyer et al: "The pAX plasmids: new gene-fusion vectors for sequencing, mutagenesis and expression of proteins in Escherichia coli", see page 129 - page 134 --	1-2,9-11
A	BIO/TECHNOLOGY, Vol. 9, 1991 George T. Davis et al: "Single chain antibody (SCA) encoding genes: one-step construction and expression in eukaryotic cells", see page 165 - page 169 --	1-11
A	NATURE, Vol. 339, 1989 Vijay K. Chaudhary et al: "A recombinant immunotoxin consisting of two antibody variable domains fused to Pseudomonas exotoxin", see page 394 - page 397 --	1-11
A	SCIENCE, Vol. 242, 1988 Robert E. Bird et al: "Single-Chain Antigen-Binding Proteins", see page 423 - page 426 --	1-11

Form PCT/ISA/

FURTHER INFORMATION CONTINUED FROM THE SEC ND SHEET

V. ☐ OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE [1]

This international search report has not been established in respect of certain claims under Article 17(2) (a) for the following reasons:

1. ☐ Claim numbers....., because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claim numbers....., because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claim numbers....., because they are dependent claims and are not drafted in accordance with the second and third sentences of PCT Rule 6.4(a).

VI. ☒ OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING [2]

This International Searching Authority found multiple inventions in this international application as follows:

See attached sheet!

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims of the international application.

2. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims of the international application for which fees were paid, specifically claims:

3. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the the claims. It is covered by claim numbers:

4. ☒ As all searchable claims could be searched without effort justifying an additional fee, the International Searching Authority did not invite payment of any additional fee.

Remark on Protest

☐ The additional search fees were accompanied by applicant's protest.

☐ No protest accompanied the payment of additional seach fees.

FURTHER INFORMATION CONTINUED

The general problem underlying the invention is not novel and a solution to it has already been found or does not involve an inventive step having regarded to the state of the art as illustrated by

a) Teeri et al. J.Cell Biocehm suppl. O (15 part G) 1991 p 217

b Markmeyer et al. Gene vol 93 (1990) p 129-134.

Therefore, the original single general inventive concept is not acceptable anymore, making it necessary to reconsider the technical relationship between the different solutions mentioned. This leads to their regrouping under distinct subjects as listed below each subject now falling under its own inventive concept.

1. Claims 5-7 completely, claims 1-4,9-11 partially. Recombinant secretable fusion proteins comprising functional protein domains joined by the hinge region of a glycosidase enzyme.

2. Claim 8 completely, claims 1-4,9-11 partially. Recombinant secretable fusion proteins comprising functional protein domains joined by the hinge region of an immunoglobulin.